

结合 Hough 变换与改进最小二乘法的直线检测

郭斯羽 翟文娟 唐 求 朱院娟

(湖南大学电气与信息工程学院 长沙 410082)

摘 要 提出了一种结合 Hough 变换与改进的最小二乘法的直线检测方法。分析了 Hough 变换与最小二乘法在直线检测及检测精度上的优缺点;利用 Hough 变换鲁棒且不需启发式信息的特点进行初步检测,确定存在直线的大致区域;利用最小二乘法确定直线区域内特征点回归直线的精确参数。为克服最小二乘法对强噪声点敏感的缺点,提出了双点移除的 p 最小二乘法,通过同时剔除具有最大正负误差的一对数据点,保证数据集中的正常点能得到可靠保留以获得精确的回归结果。实验结果表明,所提方法提高了直线检测的检测率与检测精度,且降低了对 Hough 变换的分辨率要求,可以减小算法整体的空间开销。

关键词 直线检测, Hough 变换, 最小二乘法

Combining the Hough Transform and an Improved Least Squares Method for Line Detection

GUO Si-yu ZHAI Wen-juan TANG Qiu ZHU Yuan-juan

(College of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

Abstract A novel line detection method combining the Hough transform and an improved least squares method was proposed. Advantages and drawbacks of the Hough transform and the least squares method on line detection and detection accuracy were analyzed. Robust and heuristics-free coarse detection of lines was realized by the Hough transform to determine image regions where a line may exist. Least squares regression was then applied on feature points in these regions to obtain accurate line parameters. To overcome the sensitivity of conventional least squares method to outliers, the p -least square with dual removal algorithm, which deletes each iteration a pair of data points with maximum positive and negative fitting errors to ensure the reservation of normal points in the data set and thus guarantee the accuracy of the linear regression, was proposed. Experimental results show that the novel method gives higher detection rate and more accurate line parameters compared with the Hough transform. Besides, lower Hough space resolution can be adopted without much impacts on the detection results, thus reducing the cost on the memory needed.

Keywords Line detection, Hough transform, Least squares

1 引言

直线检测是模式识别与图像理解领域基础而重要的问题。Hough 变换^[1] (Hough Transform, HT) 是解决该问题的主要方法之一, 具有鲁棒性好、无需启发式信息等优点。对 Hough 变换算法的研究主要集中在减少其计算的空间、时间开销与消除 HT 中常见的“虚假直线”问题。前者中具有代表性的方法包括 RHT^[2] 和 PPHT^[3] 等随机性方法, 此外还有采用多粒度数据融合^[4] 和预存储权值的多尺度检测^[5] 等基于标准 HT 的确定性方法; 后者通常是通过合理利用某些启发式信息来限制 HT 的累加过程或设置不同点的 Hough 空间投票权重, 如利用边缘不确定性的统计 Hough 变换^[6]、利用邻域抑制的投票加权方法^[7] 和通过线上点密度来去除部分虚假直线^[8] 等。

HT 尽管能鲁棒地检测图像中的直线, 但检测分辨率受

时间、空间及图像噪声等因素的限制。以标准直线 HT 所采用的 θ - ρ 参数空间为例, θ 和 ρ 的分辨率共同决定了 Hough 参数空间的大小, 而且 θ 的分辨率还决定了 HT 的时间性能。如果 θ 和 ρ 的分辨率取值过小, 则不仅时间上可能不能满足要求, 而且空间开销也有可能超出计算机所能承受的合理范围; 即使时间与空间的开销可以接受, ρ 的分辨率也将由于噪声造成的图像点偏离理想直线的因素而无法获得任意期望的精度, 因为过度细分的 ρ 累加器数组会由于对理想直线的偏离, 无法获得足够的累加票数而造成检测失败。

另一方面, 最小二乘法 (Least Squares, LS) 则可获得给定数据集在均方误差意义下的绝对精确直线, 从而达到 HT 所无法达到的检测精度。但是最小二乘法非常容易受所谓“野值 (outlier)”的影响, 因此需要采用改进的鲁棒回归分析方法。已有若干鲁棒回归方法, 如最小中位平方法 (Least Median of Squares, LMS)、最小截平方和法 (Least Trimmed

到稿日期: 2011-05-20 返修日期: 2011-07-31 本文受湖南大学青年教师科技创新扶持项目(531107040050)资助。

郭斯羽(1975—), 男, 博士, 副教授, CCF 会员, 主要研究方向为图像处理、机器视觉、系统建模与仿真, E-mail: syguo75@163.com; 翟文娟(1984—), 女, 硕士生, 主要研究方向为图像处理; 唐 求(1970—), 女, 副教授, 主要研究方向为虚拟仪器、智能信息处理; 朱院娟(1987—), 女, 硕士生, 主要研究方向为信号处理。

Squares, LTS)等。文献[9]较好地介绍了这些方法。但这些方法多针对大数据集和高维数据,通常需要引入重采样技术^[10],因此其计算复杂度上的优势在常见大小的图像中的直线这类低维小数据集上不一定能够得到体现。

文献[11]提出了一种利用 LS 筛选 HT 圆检测结果的方法,但 LS 对野值敏感的问题未做深入考虑。本文提出了一种改进的双点移除 p -最小二乘法(p -Least Squares with Dual Removal, p LS-DR),并与 HT 相结合来进行精确的直线检测。首先通过 HT 获得图像中存在直线的大致区域,然后对区域中的点集利用 p LS-DR 进行直线的最小二乘法检测。由于引入 p LS-DR,检测精度可明显提高,而且对 HT 的检测精度要求下降,因此可以减少 Hough 参数空间的开销以及 HT 的计算时间。

2 Hough 变换直线检测精度分析

标准直线 Hough 变换^[1]采用如下参数化直线方程:

$$x \cos \theta + y \sin \theta = \rho \quad (1)$$

式中, θ 表示直线的法线方向, $0 \leq \theta < 180$, ρ 表示原点至直线的距离(本文中 θ 的单位均为“度”, ρ 的单位均为“像素”)。通常在图像直线检测中不直接使用图像坐标系,而是使用原点在图像中心处、 y 轴方向与图像的 y 方向相反的正交坐标系,如图 1 所示的 Oxy 坐标系。

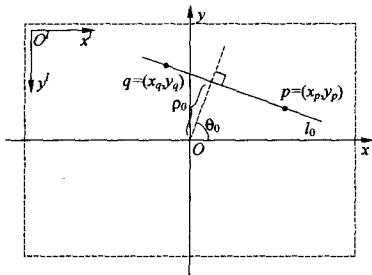


图 1 直线 Hough 变换的坐标系与参数定义

为了进行直线 HT,首先需要按一定的量化间隔将可能的 θ 与 ρ 取值范围离散化为若干区间,其中 θ 的取值范围规定在 $[0, 180)$ 的区间内,而 ρ 的取值范围则由图像矩形的顶点至原点即图像中心的距离确定;整个可能的 $\theta\rho$ 参数空间被离散化为一个二维的网格,对每一个可能的离散化参数对 (θ_i, ρ_i) 即每个网格单元设置一个计数器。然后对图像中的每个特征点 (x_0, y_0) ,遍历所有的离散 θ 值,根据式(1)计算出每个 θ_i 值下对应的 ρ 值及相应的离散区间 ρ_i ,并对计数器 (θ_i, ρ_i) 的值加 1,这一过程称为特征点对参数空间投票。当所有的特征点均完成了投票后,寻找出参数空间中计数器值大于某一给定阈值 T 的局部极大点,这些局部极大点对应的直线参数对 (θ, ρ) 即代表了 HT 检测得到的图像中的直线。

由以上 HT 直线检测的过程可见,对于给定大小的图像, Hough 参数空间的大小主要由 θ 与 ρ 的量化间隔或分辨率 $(\Delta\theta, \Delta\rho)$ 所决定。当分辨率提高到 $(\Delta\theta, \Delta\rho)/n$ 时,参数空间的大小将提高到原来的 n^2 倍;此外,由于每个特征点的投票需要遍历所有的 θ 量化单元,因此在分辨率提高到原来的 $1/n$ 时,整个投票过程的时间将增加到原来的 n 倍。所以 HT 直线检测的精度受到了时间与空间,主要是空间开销上的约束。但即使不考虑时间与空间的开销,HT 的精度也会由于图像噪声的存在无法无限提高。因为 HT 直线检测依赖于在接近

真实直线参数的离散参数单元类累积足够多的投票数,而包括数字图像量化噪声在内的图像噪声会使得特征点无法避免地偏离理想的真实直线。在较粗的 ρ 分辨率下,这样的偏离不会明显影响直线参数单元中的票数。但如果 ρ 分辨率过小,则在粗分辨率下的单元票数会被分散到精细分辨率下的若干更小的离散参数单元中,从而可能导致阈值检测失败,对后续处理造成困难。

综上所述,HT 直线检测只能达到有限精度,特别对于受噪声影响的直线,过高的精度要求可能会导致检测失败。

3 双点移除 p -最小二乘法

最小二乘法是最为常用的线性回归方法之一,能给出均方误差意义下的精确回归直线。利用 LS 处理被噪声污染的数据集是自然而直接的想法。基本的 LS 线性回归如下。

给定数据集 $\{(x_i, y_i) | 1 \leq i \leq N\}$,待确定的回归直线方程为 $y = ax + b$ 。令误差指标为

$$E = \sum_{i=1}^N (ax_i + b - y_i)^2$$

则当回归直线使得均方误差最小时,直线参数 a, b 满足

$$\partial E / \partial a = 0, \partial E / \partial b = 0$$

即

$$\sum_{i=1}^N 2(ax_i + b - y_i)x_i = 0, \sum_{i=1}^N 2(ax_i + b - y_i) = 0 \quad (2)$$

整理式(2),并令 $S_x = \sum_{i=1}^N x_i, S_y = \sum_{i=1}^N y_i, S_{xx} = \sum_{i=1}^N x_i^2, S_{xy} = \sum_{i=1}^N x_i y_i$, 则有

$$\begin{cases} S_{xx}a + S_x b - S_{xy} = 0 \\ S_x a + N b - S_y = 0 \end{cases}$$

求解方程组可得回归直线方程为

$$y = \frac{S_x S_y - N S_{xy}}{S_x^2 - N S_{xx}} x + \frac{S_y S_x - S_{xx} S_y}{S_x^2 - N S_{xx}} \quad (3)$$

由于在 HT 中采用的是如式(1)所给出的 $\theta\rho$ 参数方程,因此将式(3)重写为

$$(S_x S_y - N S_{xy})x + (N S_{xx} - S_x^2)y = S_{xx} S_y - S_{xy} S_x \quad (4)$$

对比式(1)和式(4)可得

$$\begin{cases} \theta = \arctan \frac{N S_{xx} - S_x^2}{S_x S_y - N S_{xy}} \\ \rho = \frac{S_{xx} S_y - S_{xy} S_x}{\sqrt{(N S_{xx} - S_x^2)^2 + (S_x S_y - N S_{xy})^2}} \end{cases} \quad (5)$$

式(5)即为基本 LS 所给出的回归直线参数。

但是,基本的 LS 对于明显偏离回归直线的野值很敏感,少量的强噪声就可以明显改变 LS 得到的回归直线,因此,直接利用基本 LS 来改善 HT 直线检测精度的思路并不可行。

对 LS 进行改进的一个直接想法是,根据数据点与 LS 回归直线之间的误差分布,以某个阈值将那些误差过大的数据点剔除。例如在假设数据点服从正态分布的前提下,所有与回归直线之间误差超过 2 倍或 3 倍分布标准差的数据点都可以加以去除。但实际上,这一方法通常难以奏效,这是由于强噪声的引入实际上显著提高了整个数据集的分布标准差,从而使得噪声点更容易被误认为正常数据点而无法得到剔除。

另一个方法是剔除单个误差最大的数据点,然后更新回归直线,再剔除新回归直线下误差最大的数据点,直到剩下的数据点数占原数据点数的比例为某个给定值 p 。不过这种方

法的效果仍然不佳。由式(2)的第2式可以看出,数据点与回归直线之间误差的总和为0;如果数据集中存在强度相当但误差符号相反的噪声点,则剔除某一个符号方向上的最强噪声点会造成新的回归直线更加偏向于反方向上的噪声点,这使得反方向上的噪声点更难以去除,甚至会发生正常数据点被误认为噪声点的情况。

根据以上的分析,本文提出的 $pLS-DR$ 方法在逐步剔除直至保留 p 比例数据点的思路,每一次不是仅仅去掉误差最大的单个数据点,而是同时去除正、负误差两个方向上的一对误差最大数据点,据此保证最终的回归直线主要代表误差在0附近的正常数据点。如果数据点中的强噪声点偏向于某一方向,也不会对该方法造成不良影响,因为虽然在剔除噪声点的时候牺牲了相应的正常点,但只要最终剩下的数据点主要包含的是正常点,则回归的结果不会受到大的影响。

$pLS-DR$ 的主要算法步骤如下。

给定待回归的二维点集 $P_0 = \{(x_i, y_i) | 1 \leq i \leq N\}$ 与保留数据点比例 p 。

1) 令 $i=0, n=N$ 。计算最终保留的点数 $N_f = pN$ 。

2) 利用式(5)求取 P_i 的回归直线参数 θ_i 与 ρ_i 。

3) 求取 P_i 中的每个数据点至回归直线之间的有符号误差 $e_{ij} = \rho_j - \rho_i = x_j \cos \theta_i + y_j \sin \theta_i - \rho_i, (x_j, y_j) \in P_i$ 。注意到直线方程式(1)中的 ρ 表示原点至直线之间的垂直距离,因此 $\rho_j - \rho_i$ 对应着点 (x_j, y_j) 至回归直线之间的垂直距离。这一误差度量虽然不同于 LS 中使用的 y 坐标误差,但对于确定的 θ_i ,两者之间只相差常数 $\sin \theta_i$,因此不会影响最大误差点的判断。

4) 求取分别达到最大正、负误差的两个数据点索引值 $j_{POS} = \underset{j, e_{ij} > 0}{\operatorname{argmax}} |e_{ij}|, j_{NEG} = \underset{j, e_{ij} < 0}{\operatorname{argmax}} |e_{ij}|$ 。

5) 从 P_i 中去除点 $(x_{j_{POS}}, y_{j_{POS}})$ 和 $(x_{j_{NEG}}, y_{j_{NEG}})$, 剩余的点集为 $P_{i+1}, n=n-2, i=i+1$ 。

6) 如果 $n \leq N_f$, 则返回 θ_i 与 ρ_i 作为 $pLS-DR$ 的回归结果; 否则至 2)。

图2给出了利用 $pLS-DR$ 求取回归直线的一个示例。由图可见,虽然由于数据集质量较好, $pLS-DR$ 的回归结果与 LS 的结果相差不大,但其的确有所改进,而且 $pLS-DR$ 确实有效剔除了噪声数据,从而提高了直线参数的检测精度。

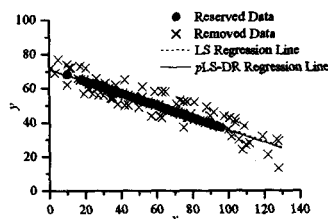


图2 LS与 $pLS-DR$ 线性回归对比

4 结合 HT 与 $pLS-DR$ 的直线检测

结合 HT 与 $pLS-DR$ 的直线检测思路是直接明了的:利用 HT 确定图像中存在直线的大致区域,然后对每个这种区域内的特征点集,利用 $pLS-DR$ 方法确定精确的直线参数。

主要步骤如下:

给定二值图像 I 、HT 的参数分辨率 $(\Delta\theta, \Delta\rho)$ 、HT 投票数阈值 T 、确定直线所在大致区域的距离误差限 d 以及 $pLS-$

DR 算法的比例参数 p 。

1) 在分辨率 $(\Delta\theta, \Delta\rho)$ 下求取 I 的标准 Hough 变换,并通过局部极大搜索来确定参数空间中投票值不小于 T 的直线参数点;

2) 对每一个检测得到的直线参数点,在 I 中寻找所有到该直线距离不大于 d 的特征点,构成特征点集 L ;

3) 利用 $pLS-DR$ 对 L 进行直线拟合,并以所得的回归直线参数作为最终检测得到的真实直线参数;

4) 重复 2)、3),直至所有 HT 检测得到的参数点都被处理完毕。

尽管提出的直线检测新方法是 HT 与 $pLS-DR$ 的一个直接的结合,但它却具有两者各自所不独立具备的优点:HT 能够在无启发式信息的情况下完成直线的粗检测,从而为 $pLS-DR$ 提供合理的数据集进行直线拟合,而且数据集的大小在一定程度上可调整,有助于减小 $pLS-DR$ 的计算负担; $pLS-DR$ 则能给出更为精确的直线检测参数,而且由于 $pLS-DR$ 本身具有一定的自适应性,因此对 HT 直线检测的精度要求也大为放宽,从而可以采用较粗的 HT 参数分辨率而减少了 HT 的时间与空间开销。

下面简要说明一下算法参数的选择。总体而言,算法对各参数的变化都不太敏感。由于 HT 只是用作直线的初步检测,因此相比于一般的 HT,新算法中的 HT 参数分辨率可以适当选取得更粗一些,以便缩小 Hough 空间的大小并减少计算时间。但过粗的分辨率会造成 HT 检测得到的直线与实际直线之间的误差增大,从而需要更大的 d 值才能保证真实直线上足够多的点落在进行 $pLS-DR$ 的点集之内,而更大的 d 值不但使 $pLS-DR$ 的运算时间延长,而且由于引入了更多的噪声点而可能影响计算精度;此外,HT 的分辨率还受图像中实际直线参数之间的差别的限制,过粗的分辨率会使得 HT 将较为接近的多条真实直线误检测为一条直线,从而造成检测准确率下降。因此,HT 的分辨率应综合以上两方面因素权衡考虑。 d 的取值则可以根据 HT 的分辨率大致估计得到。 p 的取值主要是要保证在最后剩下的点集中有足够多的正常点以保证参数估计的准确性,同时尽可能将噪声点排除掉,因此过大或过小的 p 值都将对参数估计精度产生不良影响。 T 的取值只影响 HT 检测得到的直线数量,一般取得比期望寻找的最短直线长度略小些即可。

5 实验结果与讨论

实验在一个 400 幅图片构成的图像库上进行。图片大小均为 512×512 。图片分为 8 组,每组 50 幅,各组分别加上了噪声强度 $S_n = 1\%, 3\%, \dots, 15\%$ 的均匀分布随机噪声。各幅图片中的真实直线段起止点坐标均已知。评价精度所用的误差指标是文献[7]中所使用的真实直线段至检出直线间平均距离指标 ϵ 。对某一真实直线段,找到与它之间的 ϵ 值最小的检出直线,如果这个 ϵ 值小于给定的匹配误差限 $\epsilon_{\max}$,则认为该真实直线段被确实检出,且其检测精度为 ϵ 。在实验中,设定 $\epsilon_{\max} = 1$,且 Hough 空间的峰值阈值 $T = 10$ 。

实验主要将本文算法与标准 HT 进行比较。实际上,这一比较的结果是具有代表性的,因为多数 HT 的改进主要针对降低时空开销和减少虚假直线进行,而 $pLS-DR$ 仍然可以与这些改进后的 HT 检测结果结合起来提高检测精度。

第1组实验验证算法的检测率和检测精度与 p 之间的关系。噪声强度 $S_n=1\%$, HT分辨率为 $(1,1)$, $d=2$; p 的取值为0.2, 0.3, ..., 0.6。结果如图3所示。

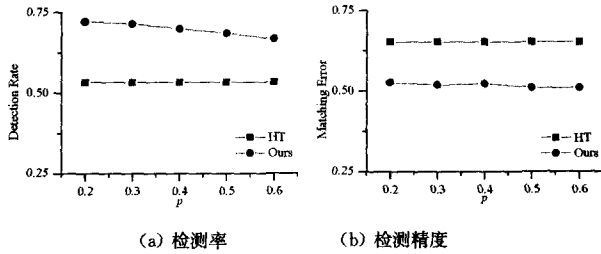


图3 直线检测率和检测精度与 p 的关系

由图3(a)可见,在给定的匹配误差限下,相比于HT,结合了 p LS-DR的算法给出了更高的检测率,这自然是 p LS-DR改善了直线参数检测精度所致,而这一点也在图3(b)中得到了体现。由图3还可以看出,算法对于 p 参数并不十分敏感,在 $p=0.2\sim 0.4$ 的较宽泛区间内,算法的检测率没有太大的差别,而检测精度所受的影响则更小。

第2组实验考查算法性能与HT分辨率之间的关系。 $\Delta\theta=0.5, 1, 2$; $\Delta\rho=0.5, 1, 2$ 。如上节所讨论过的, d 的取值在一定程度上与HT分辨率有关,因此在本组实验中,设定 $d=2\Delta\rho$; $p=0.3$; $S_n=1\%$ 。实验结果如图4所示。

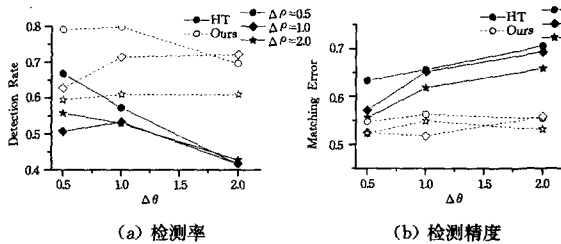


图4 直线检测率和检测精度与HT分辨率的关系

由图4(a)可见,HT分辨率的改变对HT有明显的影响,但对本文方法而言则并非如此。实际上,在当 $\Delta\theta=2$ 时, $\Delta\rho=0.5$ 的检测率比 $\Delta\rho=1$ 的检测率稍低。但在相同的分辨率下,本文方法的检测率均高于HT的检测率。图4(b)也表明在相同条件下,本文方法的检测精度明显高于HT的精度。

第3组实验考查算法与图像噪声强度之间的关系。 $S_n=1\%, 3\%, \dots, 15\%$; HT分辨率为 $(1,1)$; $d=2$; $p=0.3$ 。实验结果如图5所示。

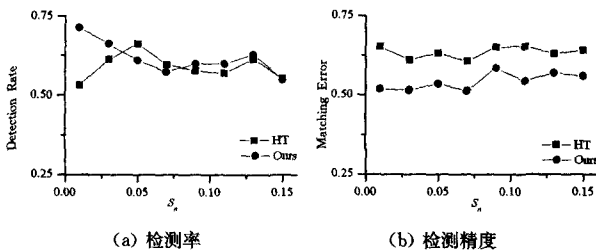


图5 直线检测率和检测精度与图像噪声强度 S_n 的关系

图5(a)表明,随着图像噪声的加强, p LS-DR算法遭遇到的噪声点增多,从而影响了该算法的性能。因此在较高的噪声下,本文算法的检测率与HT的检测率没有明显差别。但是图5(b)表明,本文算法所检测出的直线,其精度要优于HT的检测结果。

最后一组实验考查本文算法的时间性能。实验中HT的分辨率均为 $(1,1)$ 。图6(a)是 $d=2$, $S_n=1\%$ 时运行时间与 p 的关系;图6(b)是 $p=0.3$, $S_n=1\%$ 时运行时间与 d 的关系;图6(c)是 $p=0.3$, $d=2$ 时运行时间与 S_n 的关系。

由图6可见,增加 p 、减小 d 或减小 S_n 均可减少本文算法的计算时间,因为增加 p 相当于减少了 p LS-DR算法要剔除的数据点个数,从而减少了循环的次数;减小 d 或减小 S_n 均能减少 p LS-DR要处理的数据集大小;而且减小 S_n 还能减少图像中特征点的数目,使得搜索直线区域内特征点的步骤能更快完成。由图6还能发现,真正执行直线拟合任务的 p LS-DR步骤所用的时间通常都比HT检测完成后的整个后处理步骤的时间小1个数量级左右,说明搜索直线区域内特征点的步骤成为了瓶颈,对此步骤进行优化将显著提高算法的速度。图6(c)表明, p LS-DR算法和总的直线检测算法的时间复杂度阶次较高。实际上,不难看出两者的时间复杂度均为 $O(n^2)$, n 为需要处理的特征点个数。因此,对于特别大的图像和过大的 d 值,本文算法的运行时间将明显增加,此时需要采用文献[10]中给出的重采样等技术来加以处理。

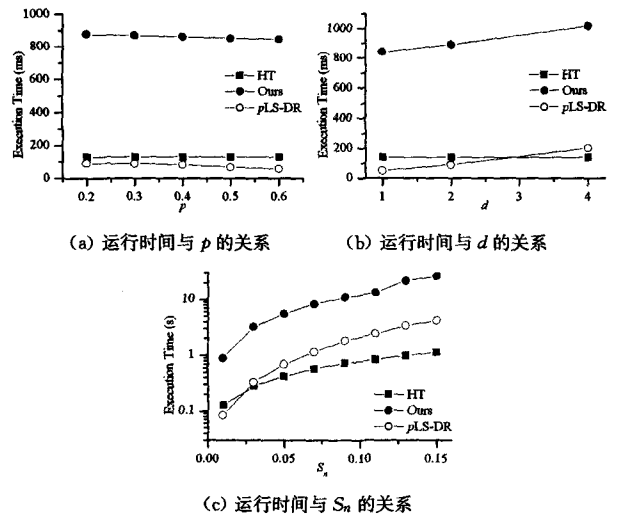


图6 算法运行时间与不同参数的关系

结束语 通过分析HT与LS在直线检测精度上的优缺点,提出了一种结合两者的直线检测方法。首先利用HT进行直线大致区域的检测,然后利用LS对该区域中的特征点进行回归分析以得到精确的直线参数。为了克服LS对野值敏感的缺点,提出了双点移除的 p 最小二乘法 p LS-DR。实验结果表明了本文算法的有效性。

目前本文方法的主要问题在于:直线区域内特征点的搜索未进行优化,极大影响了算法的运行速度;此外, p LS-DR算法的时间复杂度偏高,这虽然对一般大小的图像影响不大,但仍然需要进一步研究算法失效的临界点,以及超出临界点后的处理方法。

参考文献

- [1] Duda R O, Hart P E. Use of the Hough transform to detect lines and curves in pictures [J]. Communications of the ACM, 1972, 15(1): 11-15
- [2] Xu L, Oja E. Randomized Hough Transform (RHT): Basic mechanisms, algorithms, and computational complexities [J].

- [3] Matas J. Robust detection of line using the progressive probabilistic hough transform [J]. Computer Vision and Image Understanding, 2000, 78(1):119-137
 - [4] 谢维波,王永初,郑艺泉. 基于多粒度数据融合的直线检测算法[J]. 计算机科学, 2007, 34(9):213-217
 - [5] 徐胜华,朱庆,刘纪平,等. 基于预存储权值矩阵的多尺度 Hough 变换直线提取算法[J]. 测绘学报, 2008, 37(1):83-88
 - [6] Dahyot R. Statistical Hough transform [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2009, 31(8):1502-1509
 - [7] Guo S Y, Prodmore T, Kong Y G, et al. Improved Hough transform voting scheme utilizing surround suppression [J]. Pattern Recognition Letters, 2009, 30(13):1241-1252
 - [8] 魏志强,孙亚兵,纪筱鹏,等. 数字城市中矩形建筑物区域的自动获取[J]. 计算机科学, 2009, 36(1):211-215
 - [9] Hubert M, Rousseeuw P J, van Aelst S. High-breakdown robust multivariate methods [J]. Statistical Science, 2008, 23(1):92-119
 - [10] Hawkins D M, Olive D J. Inconsistency of resampling algorithms for high breakdown regression estimators and a new algorithm [J]. Journal of the American Statistical Association, 2002, 97(457):136-159
 - [11] 张祥德,王琪,黄亚平,等. 采用最小均方误差筛选参数的 Hough 变换及其应用[J]. 东北大学学报:自然科学版, 2007, 28(9):1270-1273
-
- (上接第 171 页)
- [11] 刘向宇,杨晓春,于戈. 一种基于特征类的高精度隐私保护数据发布方法[J]. 计算机科学, 2005, 32(7A):368-373
 - [12] 杨晓春,刘向宇,王斌,等. 支持多约束的 K-匿名化方法[J]. 软件学报, 2006, 17(5):1222-1231
 - [13] Lefevre K, Dewitt D, Ramakrishnan R. Mondrian multidimensional k-anonymity [C]//Proceedings of the 22nd International Conference on Data Engineering. 2006:25-34
 - [14] Machanavajjhala A, Gehrke J, Kifer D. L-diversity: Privacy beyond k-anonymity [C]//Proceedings of the 22nd IEEE International Conference on Data Engineering(ICDE'06). 2006:24-36
 - [15] Li Jiu-yong, Wong R C W, Fu A C, et al. Achieving k-anonymity by clustering in attribute hierarchical structures [C]//Proceedings of the 8th International Conference on Data Warehousing and Knowledge Discovery. 2006:405-416
 - [16] Wong R C, Li J, Fu A W, et al. (a, k)-anonymity: an enhanced k-anonymity model for privacy-preserving data publishing [C]//Proceedings of 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2006:754-759
 - [17] Byun J-W, Kanira A, Bertino E, et al. Efficient k-anonymity using clustering technique [R]. CERIAS Technical Report 2006-10. West Lafayette, Indiana, USA: Purdue University, 2006
 - [18] Truta T M, Vinay B. Privacy protection: p-Sensitive k-anonymity property [C]//Proceedings of the 22nd International Conference on Data Engineering Workshops. 2006:94
 - [19] Aggarwal G, Feder T, Kenihapadi K, et al. Achieving anonymity via clustering [C]//Proceedings of the 25th ACM SIGMOD SIGACT SIGART Symposium on Principles of Database Systems. 2006:153-162
 - [20] Li N, Li T. t-closeness: privacy beyond k-anonymity and l-diversity [C]//Proceedings of IEEE 23rd International Conference on Data Engineering. 2007:106-115
 - [21] 王智慧,许俭,汪卫,等. 一种基于聚类的数据匿名方法[J]. 软件学报, 2010, 21(4):680-693
 - [22] 韩建民,岑婷婷,虞慧群. 数据表 K-匿名化的微聚集算法研究[J]. 电子学报, 2008, 36(10):2021-2029
 - [23] 董云海,陶有东,唐世渭,等. 隐私保护数据发布中身份保持的匿名方法[J]. 软件学报, 2010, 21(4):771-781
 - [24] Xiao Xiao-kui, Tao Yu-fei. Anatomy: Simple and effective privacy preservation [C]//Proceedings of the 32nd International Conference on Very Large Databases. 2006:139-150
 - [25] 杨晓春,王雅哲,王斌,等. 数据发布中面向多敏感属性的隐私保护方法[J]. 计算机学报, 2008, 31(4):574-586
 - [26] 刘喻,吕大鹏,冯建华,等. 数据发布中的匿名化技术研究综述[J]. 计算机应用, 2007, 27(10):2361-2364
 - [27] 王平水,王建东. 匿名化隐私保护技术研究综述[J]. 小型微型计算机系统, 2011, 32(2):248-252
 - [28] 曾春,邢春晓,周立柱. 个性化服务技术综述[J]. 软件学报, 2002, 13(10):1952-1960
 - [29] Xiao Xiao-kui, Tao Yu-fei. Personalized privacy preservation [C]//Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data. New York, NY, USA, 2006:229-240
 - [30] Ye X J, Zhang Y W, Liu M. A Personalized (a, k)-Anonymity Model [C]//Proceedings of the 9th International Conference on Web-Age Information Management(WAIM'08). 2008:341-348
 - [31] 刘明,叶晓俊. 个性化 K-匿名模型[J]. 计算机工程与设计, 2008, 29(2):282-286
 - [32] Shen Y G, Liu Y H, Zhang Y L. Personalized-granular k-Anonymity [C]//Proceedings of International Conference on Information Engineering and Computer Science(ICIECS'09). 2009:1-4
 - [33] 申艳光,刘永红. 个性化 k-匿名隐私保护方法研究[J]. 数学的实践与认识, 2010, 40(11):97-104
 - [34] 韩建民,于娟,虞慧群,等. 面向敏感值的个性化隐私保护[J]. 电子学报, 2010, 38(7):1723-1728
 - [35] Aggarwal G, Feder T. Anonymizing tables for privacy protection [EB/OL]. <http://theory.stanford.edu/rajeev/privacy.html>, 2004
 - [36] Meyerson A, Williams R. On the complexity of optimal k-anonymity [C]//Proceedings of the ACM SIGMOD-SIGACT-SIGART Principles of Database Systems. ACM Press, 2004:223-228